university of groningen

The Role of Geographical Proximity in University-University and University-Industry Collaborations

A quantitative research focused on the city of Groningen

Author: Else Hofstra (S4106822)

Supervisor: Milad Abbasiharofteh

University of Groningen

Faculty of Spatial Sciences

BSC Human Geography and Planning

June 2022

ABSTRACT

This research focuses on the role of geographical proximity in forming collaborative ties between different universities (U-U collaboration) and between a university and industry (U-I collaboration). These collaborative ties are important in creating innovation and innovation is a major predictor of economic growth. So, it is important to gather more knowledge on collaborative ties to shape policies in a way that stimulates innovation. The main research question focuses on to what extent there is a difference in the role of geographical proximity between these two types of collaborations. The aim of the research is thus to find out whether U-U collaborations and U-I collaborations have a different relation to geographical proximity. The research focuses on the University of Groningen, located in the north of the Netherlands. A quantitative method is used to address this question. Secondary data from a patent database is processed and manually prepared for statistical data analysis via a t-test and Mann-Whitney test. The results indicate that geographical proximity is important in explaining both U-U and U-I collaborations. However, geographical proximity is more important for U-I collaborations as this combination of different organizations makes use of more tacit knowledge and deals with different incentive structures and institutions. The results recommend regional policies to focus even more on the promotion of U-I collaboration, as U-I collaboration especially takes place at a regional scale. Governments could encourage industries to locate their companies at university campuses. Future research should dive into the joint effect of multiple proximity dimensions on U-I collaborations. Moreover, future work could include the values of patent applications, instead of only counting patents.

Key terms: Geographical proximity, Innovation, Collaborative ties, Institutional proximity, Patents, Universities, and Industries.

TABLE OF CONTENTS

1. INTRODUCTION

As contemporary society becomes more knowledge-based, innovation is everywhere. Increasingly more firms see innovation as their goal and even create innovation managers within their firm (Kahn, 2018). In determining the amount of innovation, an increase in collaboration with external organisations is found to be an important factor (O'Connor et al., 2020). As an illustration, Crescenzi et al. state; 'more than 80% of all patents are registered to more than one inventor, suggesting that collaboration in research and innovation has become the norm' (p. 177, 2013). Firms can form collaborative ties with other firms but also with universities, consumers, and consultants. However, not all external actors can provide equally relevant knowledge for innovation (O'Connor et al., 2020).

As knowledge becomes more important for economic growth, academia have become key actors in the innovation system (Göransson & Brundenius, 2010; Petruzzelli, 2011). Since the digital revolution scientists across different universities began to share common resources (Cummings & Kiesler, 2007). This resulted in collaborations between universities from various continents (now; U-U collaborations). Universities can complement each other in their unique form of expertise. In addition, policy changes have encouraged scientists and industry to cooperate to develop innovative technologies (Cummings & Kiesler, 2007). Therefore, collaborations between universities and industry (now; U-I collaborations) are increasing in importance, especially in science-based fields (Ponds et al., 2007). Universities link more directly to industries through e.g., employment by firms of university students, joint research programmes or publications, and granting licenses for university patents (Petruzzelli, 2011). This causes a discussion on the appropriate role for universities in the innovation system. Should the university focus on applied research for industries or stay with generic research and teaching tasks? Either way, many universities worldwide are undergoing this process of a changing purpose (Etzkowitz & Leydesdorff, 2000). Etzkowitz and Leydesdorff (2000) are in favour of this so-called 'third mission' of universities, they should also focus on applied research in the innovation system. Since it has been found that it is more productive and profitable to combine both activities (Etzkowitz & Leydesdorff, 2000). There are also critiques of the new role of universities. The incentive system of universities seems not to be in line with that of industries (O'Connor et al., 2020). Universities strive for disclosure; they want to publicise their findings to the scientific world. On the other hand, firms reward the exploitation of knowledge by applying for patents. Therefore, you can argue that universities and firms are institutionally distant. However, U-I collaborations are forming nevertheless. In particular, Blumenthal et al. executed a national survey of universities in the USA on data withholding (1997). They found that approximately 20% of the respondents delayed the publication of their research by 6 months because this gave room for patent application (Blumenthal, 1997). In these cases, U-I collaborations were thus realized. An important factor that can solve this institutional distance and thus enhance U-I collaborations is geographical proximity. The tacit nature of knowledge requires some face-to-face contact and geographical proximity can grant this (Ponds et al., 2007). In addition, knowledge spillovers from universities are spatially bordered, so other organizations need to be closely located to benefit from this knowledge (Ponds et al., 2007). However, there is a discussion on the importance of geographical proximity. Some scholars argue that geographical proximity is not necessary for forming collaborative ties, but other forms of proximity (social, institutional, organizational & cognitive) can take on this role (Ponds et al., 2007).

Even though there are some societal concerns about U-I collaborations, government policy that encourages them is one of the most used means to stimulate innovation (D'Este, & Iammarino, 2010). For example, the European Commission already initiated policies to stimulate U-I collaboration (Petruzelli, 2011). So firstly, this paper will look into the general effect of geographical proximity on collaborative tie forming. As collaborative ties are an important predictor of innovation and geographical proximity can provide face-to-face contact that can enhance knowledge exchange. Secondly, it is

important to gather more knowledge on specifically U-I collaborations to shape policies in the right way. Therefore, this paper will dive into the effect of geographical proximity on forming collaborative ties between universities and firms. This provides knowledge on how industries form collaborations with universities. For example, if geographical proximity is important in determining collaborations, governments might invest in their regional or national scientific infrastructure (Arundel & Geuna, 2004). To make this process clearer, this paper will compare the collaborative tie forming of U-I and U-U collaborations to search for their differences. This is of relevance to policy decisions about the level of encouragement of U-I collaborations and about the fields of science that deserve this support (Arundel & Geuna, 2004). Academically, this paper adds to the discussion on the role of geographical proximity in collaborative tie forming. It will add knowledge on how important geographical proximity still is when dealing with codified knowledge and modern communication technologies, whether it is a prerequisite or a more indirect factor.

This paper attempts to contribute to the debate on collaborative ties of universities by answering the following two research questions;

RQ1. To what extent does geographical proximity trigger collaborative tie formation?

RQ2. To what extent does geographical proximity trigger the formation of U-U and U-I collaborations differently?

To answer these questions, this paper will use multiple methods. In the upcoming section, a literature review will analyse whether geographical proximity affects collaborative tie formation. Furthermore, it will explore whether this effect differs between the formation of U-U collaborations and U-I collaborations. Then, the methodology section will discuss how the secondary data was gathered and analysed and how ethics were taken into account. More specific, the hypotheses formed along with the theoretical framework will be quantitatively tested via a t-test and a Mann-Whitney test in SPSS, using a database of patent filings. In addition. this study will focus on the collaborations in this database of which at least one of the inventors is living in the city of Groningen. Subsequently, the results section will show graphs and maps and will explain how to interpret the results of the statistical analysis. Lastly, the conclusion will summarize the main results and their implication for policies and future research.

2. THEORETICAL FRAMEWORK AND HYPOTHESES

The pioneering findings in *Principles of Economics* by Alfred Marshall tried to understand why industries allocate in specific areas (2009). Marshall found that the geographical colocation of firms increases the likelihood of collaborative tie formation among these firms (Mudambi et al., 2018). This led to concepts as agglomeration economies; the external benefits firms obtain from co-location. These concepts are thus focused on geographical proximity and how this can facilitate inter-firm collaborations. In the 1990s, the proximity literature started to develop a theoretical framework. It was found that geographical co-location is not sufficient and necessary in explaining the formation of collaborative ties (Boschma, 2005). Rather, other dimensions of proximity could be more important. Boschma distinguishes 5 types of proximities; geographical, cognitive, institutional, organizational, and social (2005). The more proximate (in any dimension) two economic actors are, the easier knowledge will be exchanged. Proximities are thus important in explaining collaborations in knowledge networks, where knowledge is transmitted between economic actors (Balland et al., 2015). The sharing of knowledge is an important predictor of innovation. So, proximities are important in explaining innovation, as they facilitate knowledge exchange to take place (Boschma, 2005).

Past empirical findings have predominantly focused on collaborations between firms. This literature review will focus on the role of geographical proximity in the formation of collaborative ties of which at least one collaborator is a university. This adds to academic knowledge as the difference between U-I collaborations and U-U collaborations and the influence of geographical proximity will be discussed. This distinction has, to my knowledge, not been explicitly researched and can therefore fill a gap in the academic literature

2.1 Geographical Proximity: Still Important despite Modern Technologies?

Geographical proximity has received the most attention in proximity research (Lazzeretti & Capone, 2016). It is defined as the absolute and relative spatial distance between two economic actors (Boschma, 2005). The co-location of organizations can cause positive externalities; knowledge spillovers, a qualified labour force, and specialised suppliers. Spatial proximity gives opportunities for repeated face-to-face contact which can lead to social capital (Lazzeretti & Capone, 2016). This facilitates the exchange of, especially, tacit knowledge. These knowledge spillovers and interactions between organizations that are due to geographical proximity, lead to innovation. However, there are different views on the role of geographical proximity. Collaborative ties can be formed through geographical proximity, but it is not always the major predictor. Social, institutional, cognitive and organizational proximity can also play big roles. As an illustration, Boschma believes geographical proximity is not a sufficient condition. The other forms of proximity may serve as substitutes to solve geographical distance (2005). In this view, geosgraphical proximity plays more of an indirect role (Ponds et al., 2007).

In science, there has been a move from individual productivity to a collaborative effort. This emergence of collaboration caused most world-leading innovations to be produced in the 21st century (Dong et al., 2017). Currently, only 15% of scientific papers are produced by only one author (Dong et al., 2017). This is predominantly due to a growing division of labour among scientists (Ponds et al., 2007). There are increasingly more specialisms and scientific subfields, forcing scientists to collaborate. Moreover, there has been a change in the scale of scientific collaborations. Since the 21st century, 20% of scientific publications are produced by scientists from different states (Dong et al., 2017). So there is a higher degree of knowledge exchange in science and access to knowledge is easier than ever (Keller & Yeaple, 2013). So, does distance still matter in collaborations? Due to the rise of the Internet, communication costs have reduced and speed, quality and barriers to interactions have improved (Von Graevenitz et al., 2021). Additionally, it has to be highlighted that knowledge is intangible and can therefore be transferred via these modern communication technologies (Keller & Yeaple, 2013). This could logically lead to a neglect of agglomeration effects, to a so-called ‘death of distance’ (Von Graevenitz et al., 2021). And thus to minor importance of the effect of geographical proximity on collaborative tie-formation.

However, the scholarly consensus still holds that knowledge exchange declines when the geographical distance increases (Von Graevenitz et al., 2021). Already in 1994, Katz found that collaborations among universities occur more frequently with collaborators in closer proximity than with collaborators at a further distance. In addition, Siegel et al. researched whether companies that were located in university science parks have higher productivity. Their results affirm this hypothesis. Moreover, using survey results, Mannsfeld and Lee find that companies rather work with local university scientists that are located within a hundred miles of the firm (1996). So, even though international collaboration has been made easier by improved transportation processes and digital communication, collaboration over greater distances continues to be more costly (Ponds, et al., 2007). Geographical proximity can facilitate face-to-face contact which can play a huge role in trying to solve the complexity and uncertainty of new research (D’Este & Iammarino, 2010). Moreover, the spatial gathering of different organizations facilitates the exchange of knowledge with a high degree of useful information (Petruzzelli, 2011). Also, the recent work of Graevenitz et al. (2021), used US trademark data to measure the spatial distribution of innovation. They found that distance has a strong, negative effect on the diffusion of innovation and

thus still emphasizes the importance of geographical proximity for collaboration and innovation. This leads to the first hypothesis;

H1; Geographical proximity is important in explaining collaborative tie formation

2.2 Geographical Proximity as a Compensator in U-I Collaborations

In this section, the paper will focus on the distinction between U-U and U-I collaborations. Therefore it is first important to specify U-U and U-I collaborations. In this paper, U-U collaborations are defined as any collaboration in which two or more universities or research institutions are mentioned as an inventor (Katz, 1994). U-I collaborations are in this paper defined as any collaboration in which one or more universities or research institutions and one or more companies are mentioned as inventors (Katz, 1994). These collaborations can be distinguished especially on their degree of institutional proximity. Organizations can be institutionally proximate in a formal way, through common laws and rules, and in an informal way, through common cultural norms and habits (Boschma, 2005). For example, a common language generates trust and can therefore simplify collaborative tie-formation. This distinction will be elaborated on in this section.

Knowledge transmission and recombination are the drivers of innovation and therefore of economic growth (Alhusen et al., 2021). Models of innovation describe innovation as an interactive process where firms collaborate with other firms, customers, government, and research institutions (Jensen et al., 2007). The dominant measurements of innovation are formal R&D processes e.g., patents, publications, and expenditures (Alhusen et al., 2021). These measurements focus on formal processes of innovation and thus on codified knowledge, while these formal processes have proven to not fully explain innovation (Alhusen et al., 2021). As an alternative, Jensen et al. describe two ideal types of innovation (2007). They argue that there is a distinction in interactive learning that leads to innovative achievements. Firstly, they emphasize the difference between tacit and codified knowledge. Codified knowledge is explicit knowledge that is formally articulated and can therefore be easily transmitted, even though it is removed from its context of use (Zack, 1999). Codified knowledge can be written down, but to understand it some prior knowledge on the topic is needed (Zack, 1999). In contrast, tacit knowledge is unconsciously understood and applied and therefore more difficult to transmit. However, it can be transmitted; usually through vastly interactive communication and mutual experiences (Zack, 1999). Based on this distinction in knowledge, Jensen et al. distinguish the two modes of interactive learning that lead to innovation (2007).

The Science, Technology and Innovation (STI) mode is based on the production and adoption of codified scientific knowledge (Jensen et al., 2007). The STI mode focuses on creating generic knowledge, and to do so also retrieving knowledge from former publicised generic knowledge. Thus, STI makes use of globally accessible knowledge to potentially generate globally accessible generic knowledge (Jensen et al., 2007). The other mode of interactive learning introduced by Jensen et al. is the Doing, Using and Interacting (DUI) mode, which is based on experience-based learning and tacit knowledge (2007). The DUI mode focuses on localized knowledge which is informally transmitted between and within industries. DUI learning can be measured via problem-solving teams and job and task rotation which promote knowledge exchange (Jensen et al., 2007).

The STI-DUI dichotomy can be used in this research, as it is similar to the dichotomy between scientific research used within universities and applied innovation used within industries. The STI mode is in nature very similar to scientific research. Scientific research is focused on acquiring new knowledge, without any specific application in mind. So, it does not seek long-term economic benefits and does not aim to solve particular problems (Tijssen, 2010). Its goal is to produce codified knowledge that can be used globally. So, universities are using more of an STI mode in creating innovation. On the other hand, as the DUI mode focuses on tacit knowledge, it is similar to industrial innovation and thus related to

firms. Firms aim at applied research to find possible purposes for basic science findings or seek ways to achieve particular goals (Tijssen, 2010). In applied research, tacit knowledge is required to understand the complexity of the problem at stake.

So, an important distinction in innovation modes can be created between universities and industries. This distinction between innovation dynamics in collaborative tie forming is interesting for this paper. The different orientations of the systems impact the effect of geographical proximity on U-I and U-U collaborations. Particularly, universities make the most use of codified knowledge, which is globally transmitted and can, in most cases, be globally accessed. Even though it is more costly, scientists can collaborate conveniently over greater distances as they are used to a common codified language and an underlying incentive structure, namely global publishing (Ponds et al., 2007). There is mutual trust among universities and a common scientific language, enhancing collaboration. In other words, universities are institutionally proximate by nature and are therefore less dependent on geographical proximity. In conclusion, co-location of universities in U-U collaborations is expected to be less of a requirement in codified knowledge transfer. However, even codified knowledge requires some tacit knowledge in its interpretation, so geographical proximity does play a role (Howells, 2002).

In contrast, U-I collaborations are expected to make more use of tacit knowledge. As mentioned, to transmit tacit knowledge, intensive conversation or shared experiences are necessary. Co-location is an important requirement for such intensive knowledge sharing. The use of tacit knowledge is necessary as a U-I collaboration is complex. Different modes of innovation are used and the incentive structures differ, which influences the formal rules as well as informal norms. This makes universities and industries institutionally distant. Geographical proximity can help overcome these differences and support the establishment of successful collaborations (Ponds et al., 2007). It can enhance trust-building between companies and universities, which can make their differences superfluous. In other words, geographical proximity can step in to compensate for this lack of institutional proximity. Therefore, firms that aim to work with universities, are expected to be more dependent on geographical proximity (Arundel & Geuna, 2004). This leads to the second hypothesis;

H2; Geographical proximity triggers U-I collaborations more strongly compared to U-U collaboration

2.3 Conceptual Model

The theoretical framework and formed hypotheses lead to a conceptual model (Figure 1). This research focuses on U-U and U-I collaborations formed with at least one inventor living in the city of Groningen. It is expected that geographical proximity is important in explaining the formation of both U-U and U-I collaborative ties. Even though communication costs and efforts have decreased due to modern technologies, knowledge exchange does decline when geographical distance increases (Ponds et al., 2007). However, it is expected that geographical proximity is more important in explaining U-I collaborations, compared to U-U collaborations. Due to the higher use of tacit knowledge and the difference in institutions, firms and universities that cooperate are expected to be located in closer proximity to one another. In conclusion, both collaborations are expected to be influenced by geographical proximity, but U-I collaborations to a greater extent.

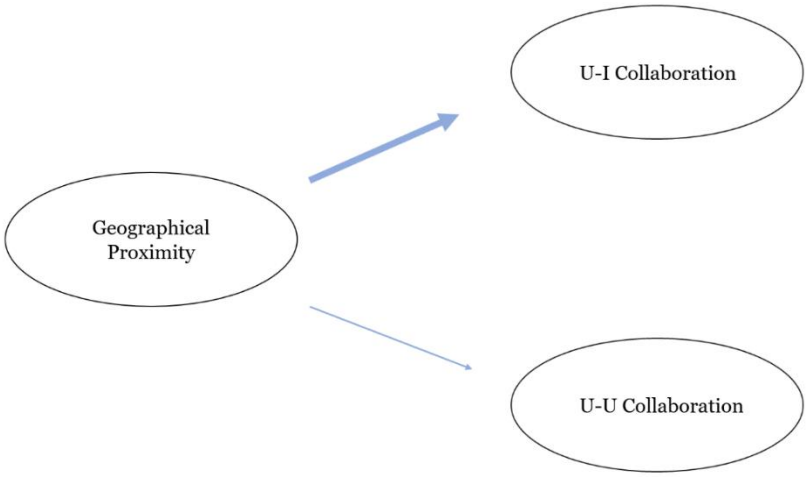

Figure 1. The conceptual model as based on the hypotheses.

3. METHODOLOGY

This paper uses quantitative methods to answer the main question. Secondary data retrieved from the OECD REGPAT database will be used and modified to analyse the relationship between geographical proximity and collaborative ties in SPSS. This study will focus on the city of Groningen. All collaborations captured in the research will thus at least have one collaborator that is living in the city of Groningen. The following section will dive deeper into the operationalization and analysis of the data and ethical rules used.

3.1 Operationalization

To answer the research questions that focus on the effect of geographical proximity on the formation of collaborative ties and the difference of this effect between U-I and U-U collaborations, statistical analysis will be used on secondary data. Secondary data seemed suitable because collecting primary data can cause time-related and sensitivity issues (Martins et al., 2018). Companies can be hesitant with sharing patent data with an unknown individual. While larger organizations often already obtained the data the researcher is looking for. So, secondary data is used to retrieve quality results at a faster pace. In this research, the secondary data is retrieved from the OECD REGPAT database, version 2021, which contains patent applications filed with multiple national and regional patent offices, such as the European Patent Office (Maraut et al., 2008). This database was suitable for this research due to its relatively small size. Additionally, the database is publicly accessible, reliable and easy to download. Moreover, the patent data is linked to regions, which enabled regional analysis of collaborative ties (Maraut et al., 2008). The database was also suitable for this analysis as it includes information on the addresses of the different collaborators. Therefore, it is possible to precisely measure geographical proximity for each collaboration that resulted in a patent application. The address information is also used to define an application as a collaboration, as each patent application which contained two or more addresses was defined as a collaboration. Altogether, it was a practical and efficient decision to make use of the OECD REGPAT database. The database was filtered by the supervisor of this research to only maintain patent filings from 2001 to 2018 of which at least one inventor's address is located within Groningen. The research is focused on this city, as the researcher is a student at the university in Groningen. Therefore, the author is interested in the innovative collaborations that the university and companies in this city participated in, in the last two decades. Table 5 in Appendix A contains

clarification on how the database is structured and how each variable will be used and processed in this research.

Each collaboration has been manually categorized as either a U-U or U-I collaboration, according to the definitions in the theoretical framework, depending on the organization addresses as captured in the database. If a collaboration was between a different pair of organizations, for example, Industry-Industry, it was treated as not useful and filtered out. The research is only interested in U-U and U-I pairs. Even though collaboration ultimately takes place between persons, the focus of this paper is on organisations (Ponds et al., 2007). All addresses have been manually checked to assure that each inventor used his/her address of employment. If not, the names of inventors were used to find out where they were employed in the year of the patent application. Subsequently, geographical proximity is measured through the spatial distance between the location sites of the universities and in some cases, companies jointly inventing (Petruzzeli, 2010). The addresses of organizations were used to calculate the distance between actors in each collaboration. This has been done via the 'calculate distance' function in Google Maps, which provides you with the absolute distance between two locations. The absolute distance is used, as contemporary communication technology facilitates collaboration without the in-person presence necessary. Moreover, the Dutch transport system is very efficient and when collaborators want to meet in person, the relative distance is strongly correlated to the absolute distance. In the case of a patent filing with more than two collaborators, for each link, the geographical distance is measured. Then, the average was used to produce the geographical proximity variable for this collaboration. The LOG of each geographical distance in kilometres will be taken to minimize the strong effect of outliers. In conclusion, this paper utilizes quite some manual check and filter work, which is time-consuming. This also explains the relatively small number of patent filings analysed.

3.2 Data Analysis

Hypothesis 1

The first expectation is that geographical proximity is important for collaborative tie-formation. This is analysed by using the data of the actual U- U and U-I collaborations as captured in the REGPAT database and comparing it with a null expectation. The null expectation is measured by creating random pairs among the filtered organizations in the database and then measuring the absolute distance between those organizations. So, the organizations that formed U-U and U-I collaborations in the database are used to create random pairs. Each organisation gets assigned a number which belongs to another organisation in the database, forming a collaborative tie. So, this resulted in randomly created collaborations, consisting of two organizations. Subsequently, the absolute distance between the organisations within these collaborations was calculated. This resulted in the geographical proximity of the null expectation. The absolute distance is empirically used to produce the theoretical geographical proximity variable. They have a negative relation; the more absolute distance, the less geographical proximity. Via a two samples t-test was tested whether the sample means are significantly different. This way, it was analysed whether the influence of geographical proximity differs between the actual collaborations and a null expectation, namely the randomly created collaborations. The expectation is that geographical proximity is more important in the actual collaboration than in the null expectation. It will ensure that the influence of geographical proximity is not a coincidence. This way, I can prove or disprove hypothesis 1 and answer the first research question.

Hypothesis 2

Geographical proximity enhances the forming of U-I and U-U collaborations, as it facilitates easier face-to-face contact between collaborators. However, U-I collaborations are most dependent on tacit knowledge and are more institutionally distant, therefore proximity is expected to be more important in U-I than in U-U collaborations, as captured in Hypothesis 2. To test this hypothesis, a Mann-Whitney test is used. This test seemed suitable as we look for the difference in the variable geographical proximity between the two groups; U-U and U-I (these variables in the analysis are elaborated on in table 1). Furthermore, a Mann-Whitney test requires an ordinal or continuous variable and two independent

groups, which makes our data suitable. Lastly, the test is a nonparametric alternative for the two-samples t-test, which was necessary due to the small number of cases. The null hypothesis for this test is: *In the population, geographical proximity triggers U-I and U-U collaborations equally.* If we find that the p-value is significant, we might reject the null hypothesis and state that they are triggered differently by geographical proximity. Via the Mann-Whitney test, it can be determined whether the median of the geographical proximity variable, differs between U-U and U-I collaborations. As a result, it can be determined whether geographical proximity triggers the formation of U-U and U-I collaborations differently and the second research question will be answered. The author has experience in using both t-tests and Mann-Whitney tests in the Statistics and MAR courses and former research.

TABLE 1.

Variables used in the Mann-Whitney test

Variable	Indicator	Measurement Scale
Collaboration type	Addresses of the organizations as captured in the patent database. Resulting in either a U-U or a U-I collaboration	Nominal (Binary)
Geographical Proximity	The distance in kilometres between two collaborating organizations	Ratio

3.3 Research Ethics

To make a reliable research project, it is necessary to comply with ethical rules. In this paper, the Netherlands Code of Conduct for Research Integrity is used. This code consists of five aspects that need to be met by a researcher. The author applied these aspects to this research:

Honesty: The author will be open to comments and suggestions of peer-reviewers and supervisors, to improve the quality of the research. In addition, the author will be accurate about the data processing and results gained, no unfounded claims are aimed to be made.
Scrupulousness: The author will be careful and conscious of the scientific methods used in the research and of the design and reporting of the research.
Transparency: The author will be as open as possible about the data used and how this data is processed. All research steps are aimed to be verifiable and the role of external actors will be shared.
Independence: The author will operate in academic liberty and is not dependent on someone or something else. This also entails that the author will be impartial.
Responsibility: The author will keep individuals that are mentioned in the OECD REGPAT database anonymous within the paper. The addresses of inventors that are captured in the database are not shown or shared in this research. The names of the inventors that are captured in the database and are used to have further information on their employment are only used for research purposes. The documents used to produce this research are only accessible to the author to further secure anonymity. (Netherlands Code of Conduct for Research Integrity, 2018).

4. RESULTS

4.1 Geographical Proximity and Collaborative Tie-Formation

In order to affirm or reject Hypothesis 1 (*Geographical proximity is important in explaining collaborative tie formation*), the logs of the absolute distance were compared across the actual collaborations and randomly assigned collaborations. A two samples t-test was used to compare the

sample means. However, a two-sample t-test requires a normal distribution in both groups. Therefore, a Shapiro-Wilk test was used to test the normality of the distributions. The null hypothesis of this test was; *the DistanceLOG variable is normally distributed in both populations; Actual and Random*. The results of the Shapiro-Wilk showed that we can significantly say that the distribution is normal in only one of the populations; Actual collaborations; W(36)= .92, p= 0.01 and Random collaborations; W(42)= .96, p= 0.20 (Appendix B, Table 6). Only for the distance between the random pairs we can say accept the null hypothesis that the distribution is normal. The boxplots of both populations show us that the actual collaboration population has a strong outlier and a higher interquartile range, which adds to the distribution not being normal (Appendix B, Figure 6). However, the Central Limit Theorem states that by increasing sample size, the distribution tends to be close to the normal distribution (Illowsky et al., 2013). It considers sample sizes that are above 30 in all populations sufficient to assume a normal distribution. As we have a sample size above 30 in both populations, we can assume normal distribution and pursue the two-samples t-test.

As displayed in the table below, there was a total of 36 collaborations found in the database containing at least one university and one address located in Groningen (Table 2). Within these collaborations there were 42 unique organizations, resulting in 42 randomly created pairs of two organizations. These random pairs are used as a null expectation. Already from these descriptive statistics, we can see a difference in the means; the organizations within the random pairs were located further from one another (M= 2.91, SD= 0.45) than the organizations within the actual collaborations (M=2.64, SD= 0.58).

Table 2

Descriptive statistics for hypothesis 1.

Group Statistics

	Type	N	Mean	Std. Deviation	Std. Error Mean
DistanceLOG	Actual	36	2,6443	,57637	,09606
	Random	42	2,9091	,45132	,06964

However, to see whether the difference in means is significant, an independent two samples t-test was used. The null hypothesis of this test was; *the means of the variable DistanceLOG are equal between the groups; actual collaborations and random collaborations*. The output of the t-test is displayed below (Table 3). The results show that the t-test is significant, t(76)= -2.27, p=0.026. So, the statistics show that we can reject the null hypothesis that the means across the groups are equal. And we can thus assume that there is a difference in geographical proximity between the random and actual collaborations. The t-statistic is negative and this implies that the distance between collaborators is greater within the random collaborations, as this was our second group variable (table 2). This is in line with former research, geographical proximity gathers organizations and fosters interactions with high information richness and tacit knowledge exchange (Petruzzelli, 2011). Also, in U-U collaborations using codified knowledge, the interpretation of codified knowledge still needs tacit knowledge and thus requires some geographical closeness (Howells, 2002). In conclusion, the results provide evidence to support the first hypothesis and say that geographical proximity is important in explaining collaborative tie-formation.

Table 3

The Output of the Two Samples T-Test

Independent Samples Test

		Levene's Test for Equality of Variances		t-test for Equality of Means						
									95% Confidence Interval of the Difference	
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	Lower	Upper
DistanceLOG	Equal variances assumed	,937	,336	-2,274	76	,026	-,26486	,11645	-,49680	-,03293
	Equal variances not assumed			-2,232	65,914	,029	-,26486	,11865	-,50176	-,02797

4.2 U-U and U-I Collaborations: spatially limited?

This research focuses on the city of Groningen, as each collaboration contained an inventor who was living in Groningen. When analysing the data and finding inventor's working addresses it became clear that each collaboration contained an employee of the University of Groningen. Therefore, the research now focuses specifically on collaborations of the University of Groningen with other universities or companies. The university is spread out across the city of Groningen and thus makes use of different buildings. The map in Figure 2 shows that three of these locations were captured in the OECD REGPAT database; The Zernike campus, the 'Academiegebouw' (Academy building), and the University Medical Centre Groningen (UMCG). Most patent applications were registered at the address of the UMCG.

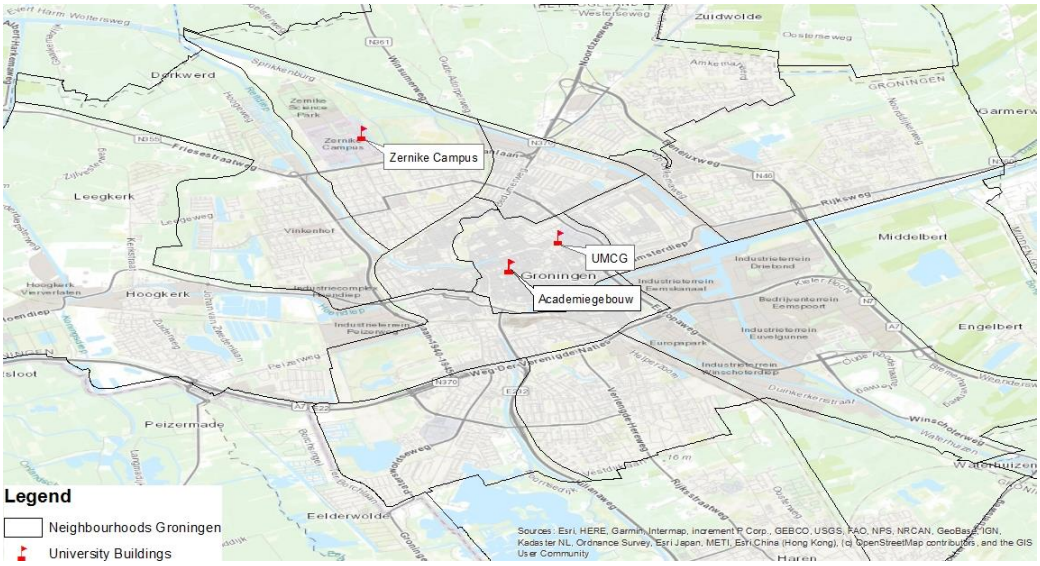

Figure 2. Map displaying the locations of the University of Groningen buildings captured in the database. Created by the author (2022).

After the full data processing and filtering process for the distinguishment between U-U and U-I collaborations, the descriptive results were as displayed in the table below. As visible in table 4, in the database, I found 20 U-I collaborations between different organizations and 16 U-U collaborations between different organizations. These collaborations are displayed on the map in Figure 3. As shown, almost all collaborations with the University of Groningen were in Europe, again emphasizing the importance of geographical proximity. Only two organizations were located on a different continent, all in the United States.

Table 4

Descriptive Statistics Hypothesis 2

Type of Collaboration

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	U-I	20	55,6	55,6	55,6
	U-U	16	44,4	44,4	100,0
	Total	36	100,0	100,0	

Figure 3. Map displaying the overview of all captured U-U and U-I collaborations between organizations in the database. Created by the author (2022).

As already mentioned in the methodology section, the Mann-Whitney test is suitable for the analysis. This test sorts all values from low to high and then attributes rank numbers to each value; 1 to the lowest value and so forth. This makes it possible to test whether there is a significant difference in the mean ranks. The null hypothesis of the Mann-Whitney test was; *In the population, the mean ranks of the logs of the average distance between collaborators in the two groups, U-U and U-I, are equal*. The results show that we can reject this null hypothesis, p=.000 (Figure 4). This tells us that there is a significant difference between the mean ranks of distance in U-U and U-I collaborations. Furthermore, the results provide information on the distributions of the mean ranks in both groups. The mean rank of the distance between collaborators is greater in U-U collaborations (MR=25.38) than in U-I collaborations (MR=13.00) (Figure 4). This entails that a firm and a university from which inventors collaborate, tend to be located closer to one another than when universities or research centres collaborate. In other words, geographical proximity seems to be more important in U-I collaborations compared to U-U collaborations. This conclusion was expected when reading academic literature. Universities use codified knowledge and an STI mode in creating innovation (Jensen et al, 2007). U-U collaborations deal with two of the same organization types and therefore the same institutions. This decreases their dependence on geographical proximity. On the other hand, U-I collaborations deal with two different organization types, with contrasting institutions and innovation modes. This increases the necessity of tacit knowledge and therefore makes geographical proximity more important (Arundel & Geuna, 2004). This conclusion was also visible from the map in Figure 5. This map displays all the collaborations

between universities and universities and industries within Europe captured in this research. It stands out that the U-I collaborations in red are all clustered around the city of Groningen, while the U-U collaborations in blue spread out more across Europe. The statistical test and visualization lead to the same conclusion; the second hypothesis can also be confirmed.

Independent-Samples Mann-Whitney U Test Summary

Total N	36
Mann-Whitney U	270,000
Wilcoxon W	406,000
Test Statistic	270,000
Standard Error	31,391
Standardized Test Statistic	3,504
Asymptotic Sig.(2-sided test)	,000
Exact Sig.(2-sided test)	,000

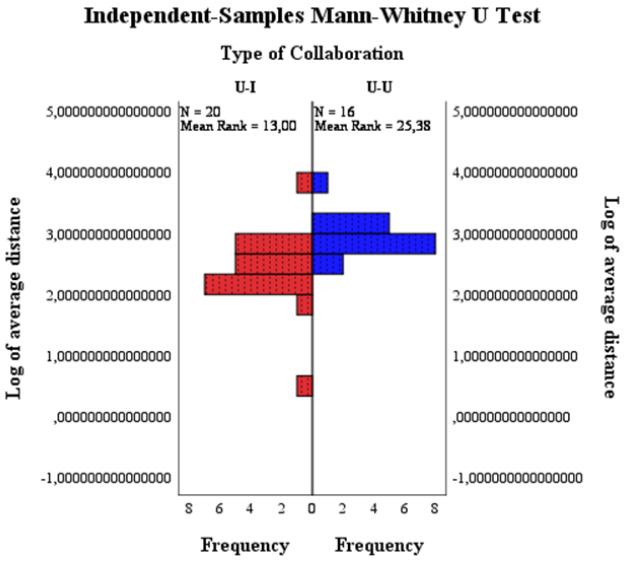

Figure 4. Results of the Mann-Whitney test. On the left, is the probability table and on the right, a distribution graph containing the grouping variables; U-U and U-I. Created by the author (2022).

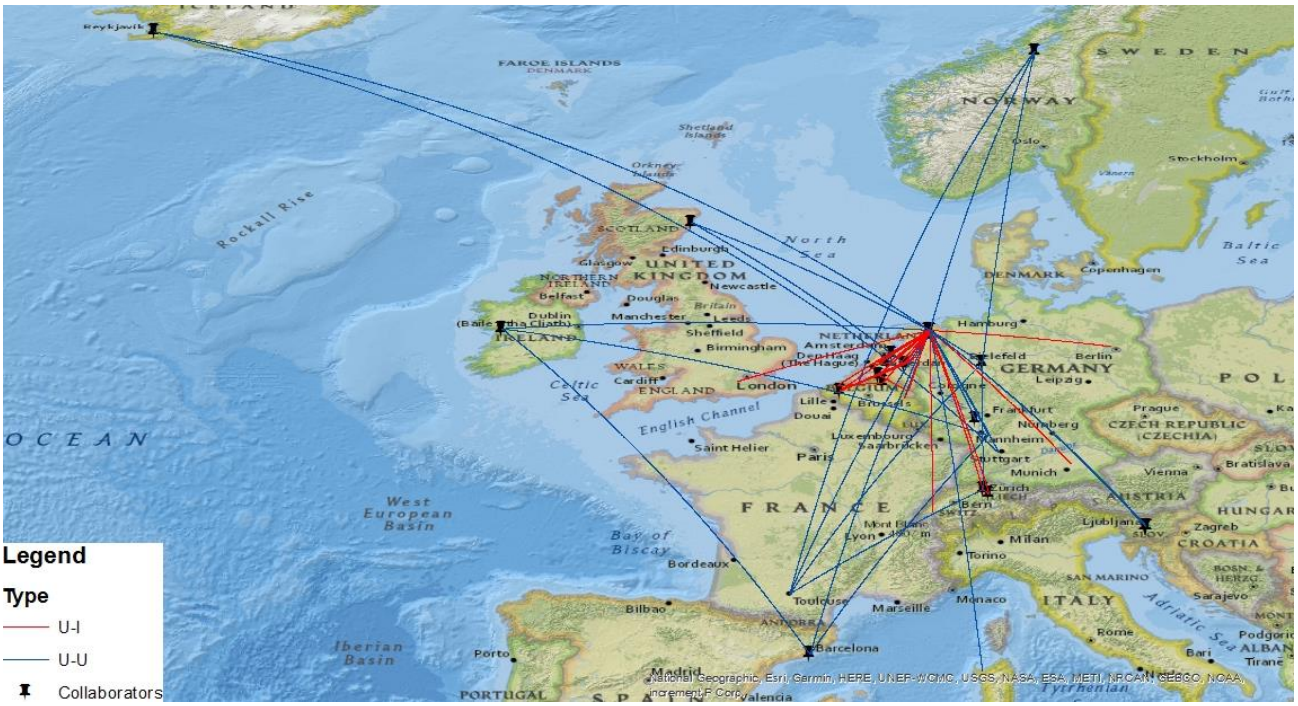

Figure 5. Map displaying the U-U and U-I collaborations in Europe. Created by the author (2022).

5. CONCLUSION

Collaboration with universities is important in creating innovation, universities can share their expert knowledge with other universities or firms. They can then form collaborative ties, resulting in U-U and U-I collaborations. This research dived into the following questions; *To what extent does geographical proximity trigger collaborative tie formation?* & *To what extent does geographical proximity trigger the formation of U-U and U-I collaborations differently?*

The findings contribute to the literature by supporting previous empirical results. This research has shown that both hypotheses following from the theoretical framework can be confirmed. Geographical proximity is important in explaining collaborative tie formation. Geographical proximity may facilitate face-to-face contact which can be of great importance when dealing with complex and uncertain new research (D'Este & Iammarino, 2010). And even though digital communication has made international collaboration easier, it is still more costly (Ponds et al., 2007). Furthermore, based on this research, we can say that geographical proximity does trigger the formation of U-U and U-I collaborations differently. That is to say, geographical proximity is more important in explaining U-I collaboration, than U-U collaboration. Universities use an STI mode in creating innovation, they use codified knowledge and have the same incentive structures. U-I collaborators use different modes in creating innovation and therefore make use of tacit knowledge which is more difficult to transfer. Tacit knowledge and the difference in incentive structures between universities and firms complexify collaboration. Geographical proximity can solve this complexity, by enabling intensive conversations or shared experiences. In line with Ponds et al., who concluded that geographical proximity is more important in collaborations between an academic and a non-academic organisation than in purely academic collaboration (2007). Boschma (2005) mentioned in his research that geographical proximity plays more of an indirect role in creating collaborative ties, but this research gives geographical proximity a more prominent role. It is plausible that geographical proximity can compensate for the lack of other proximities, such as in this research; institutional proximity.

A limitation of this study is the small number of cases that could be drawn from the database. Much manual work was devoted to data processing and analysis, so the number of cases could not be much higher as this would make the manual work unattainable. The low number of cases also led to the use of a nonparametric test, which decreases the explanatory strength of the results. So, for future research, it is recommended to examine U-I collaborations more precisely by using a higher number of cases in the analysis. Moreover, future research could focus on the joint effect of the proximity dimensions on U-I collaborations. In addition, this study counts the number of patents but does not look into the value of patents created (Siegel et al., 2003). Future research could add the value of patents by for example looking at citations related to the patent created. This will add to the knowledge about the importance of geographical proximity in U-I collaborations.

Government policies that try to stimulate innovation are already focused on encouraging U-I collaboration. Geographical proximity is quite important in the forming of U-I collaborative ties. Collaborative ties are very important in creating innovation and innovation is one of the main drivers of economic growth. Based on the results, governments could focus more on improving their regional scientific infrastructure to simplify the forming of these collaborative ties and possibly increase their economic growth. They can for example encourage firms to locate at university campuses or science parks to stimulate U-I collaboration. After all, firms located at science parks have higher research productivity (Siegel et al., 2003). In addition, policy designs that want to promote collaboration over larger distances should take the importance of geographical proximity into account.

REFERENCES

Alhusen, H., Bennat, T., Bizer, K., Cantner, U., Horstmann, E., Kalthaus, M., Proeger, T., Sternberg, R., & Töpfer, S. (2021). A New Measurement Conception for the 'Doing-Using-Interacting' Mode of Innovation. *Research Policy*, *50*(4), 104214. https://doi.org/10.1016/j.respol.2021.104214

Arundel, A., & Geuna, A. (2004). Proximity and the use of public science by innovative European firms. *Economics of Innovation and new Technology*, 13(6), 559-580.

Balland, P. A., Boschma, R., & Frenken, K. (2015). Proximity and innovation: From statics to dynamics. *Regional Studies*, *49*(6), 907-920.

Benner, M., & Sandström, U. (2000). Institutionalizing the triple helix: research funding and norms in the academic system. *Research policy*, *29*(2), 291-301.

Blumenthal, D., Campbell, E. G., Anderson, M. S., Causino, N., & Louis, K. S. (1997). Withholding research results in academic life science: evidence from a national survey of faculty. *JAMA*, 277(15), 1224-1228.

Boschma, R. (2005). Proximity and innovation: a critical assessment. *Regional studies*, *39*(1), 61-74.

Crescenzi, R., Nathan, M., & Rodríguez-Pose, A. (2013). Do Inventors Talk to Strangers? On Proximity and Collaborative Knowledge Creation. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.2367672

Cummings, J. N., & Kiesler, S. (2007). Coordination costs and project outcomes in multi-university collaborations. *Research policy*, *36*(10), 1620-1634.

D'Este, P., & Iammarino, S. (2010). The spatial profile of university-business research partnerships. *Papers in regional science*, 89(2), 335-350.

Dong, Y., Ma, H., Shen, Z., & Wang, K. (2017). A century of science: Globalization of scientific collaborations, citations, and innovations. In *Proceedings of the 23rd ACM SIGKDD international conference on knowledge discovery and data mining* (pp. 1437-1446).

Etzkowitz, H., & Leydesdorff, L. (2000). The dynamics of innovation: from National Systems and "Mode 2" to a Triple Helix of university–industry–government relations. Research Policy, 29(2), 109–123. https://doi.org/10.1016/s0048-7333(99)00055-4

Göransson, B., & Brundenius, C. (2010). *Universities in Transition*. Springer Publishing. https://doi.org/10.1007/978-1-4419-7509-6

Howells, J. R. (2002). Tacit knowledge, innovation and economic geography. *Urban studies*, *39*(5-6), 871-884.

Illowsky, B., & Dean, S. (2013). *Introductory Statistics*. OpenStax. https://openstax.org/books/introductory-statistics/pages/1-introduction

Jensen, M. B., Johnson, B., Lorenz, E., & Lundvall, B. K. (2007). Forms of knowledge and modes of innovation. *Research Policy*, *36*(5), 680–693. https://doi.org/10.1016/j.respol.2007.01.006

Kahn, K. B. (2018). Understanding innovation. *Business Horizons*, *61*(3), 453-460.

Katz, J. (1994). Geographical proximity and scientific collaboration. *Scientometrics*, *31*(1), 31-43.

Keller, W., & Yeaple, S. R. (2013). The Gravity of Knowledge. *American Economic Review*, *103*(4), 1414–1444. https://doi.org/10.1257/aer.103.4.1414

Lazzeretti, L., & Capone, F. (2016). How proximity matters in innovation networks dynamics along the cluster evolution. A study of the high technology applied to cultural goods. *Journal of Business Research*, *69*(12), 5855–5865. https://doi.org/10.1016/j.jbusres.2016.04.068

Mannsfeld, E., & Lee, J. Y. (1996). The modern university: contributor to industrial innovation and recipient of industrial R&D support. *Research policy*, *25*(7), 1047-1058.

Maraut, S., Dernis, H., Webb, C., Spiezia, V., & Guellec, D. (2008). The OECD REGPAT database: A Presentation. *OECD Science, Technology and Industry Working Papers*.

Marshall, A. (2009). *Principles of economics: unabridged eighth edition*. Cosimo, Inc..

Martins, F. S., Cunha, J. A. C. D., & Serra, F. A. R. (2018). Secondary Data in Research – Uses and Opportunities. *Revista Ibero-Americana de Estratégia*, *17*(04), 01–04. https://doi.org/10.5585/ijsm.v17i4.2723

Mudambi, R., Narula, R., & Santangelo, G. D. (2018). Location, collocation and innovation by multinational enterprises: A research agenda. *Industry and innovation*, *25*(3), 229-241.

Netherlands Code of Conduct for Research Integrity (2018).

O'Connor, M., Doran, J., & McCarthy, N. (2020). Cognitive proximity and innovation performance: are collaborators equal? *European Journal of Innovation Management*, *24*(3), 637–654. https://doi.org/10.1108/ejim-11-2019-0347

OECD, REGPAT database, January 2021.

Petruzzelli, A. M. (2011). The impact of technological relatedness, prior ties, and geographical distance on university–industry collaborations: A joint-patent analysis. *Technovation*, *31*(7), 309–319. https://doi.org/10.1016/j.technovation.2011.01.008

Ponds, R., Van Oort, F., & Frenken, K. (2007). The geographical and institutional proximity of research collaboration. *Papers in regional science*, *86*(3), 423-443.

Siegel, D. S., Westhead, P., & Wright, M. (2003). Assessing the impact of university science parks on research productivity: exploratory firm-level evidence from the United Kingdom. *International Journal of Industrial Organization*, *21*(9), 1357–1369. https://doi.org/10.1016/s0167-7187(03)00086-9

Tijssen, R. J. (2010). Discarding the 'basic science/applied science'dichotomy: A knowledge utilization triangle classification system of research journals. *Journal of the American society for Information science and Technology*, *61*(9), 1842-1852.

Von Graevenitz, G., Graham, S. J. H., & Myers, A. F. (2021). Distance (still) hampers diffusion of innovations. *Regional Studies*, *56*(2), 227–241. https://doi.org/10.1080/00343404.2021.1918334

Zack, M. H. (1999). Managing codified knowledge. *Sloan management review*, *40*(4), 45-58.

APPENDIX

A. OECD REPGAT Database Variables.

TABLE 5

Variables in the OECD REGPAT Database.

Variable as in the database	Meaning	Use for this Research
app_nbr	EPO application number	Used to identify the unique patent filings to find all collaborators
person_id	Unique person identifier	Used to identify all unique collaborators within one patent filing
address	Address of the inventor	Used to measure the geographical distance between collaborators and decide whether a filing can be defined as a collaboration
inv_name	Name of the inventor	Used to check whether collaborators filled in their home or work address. It is important to ensure that only work addresses are included in the analysis. Using the names of collaborators, it can be checked via Google where they are employed, and thus whether they used their work address.
inv_share	Inventors' share	Used as additional information on how many collaborators there are involved in one patent application
prio_year	Year of the first application	Used to make sure only patent applications from 2001-2018 are included. The analysis is limited to these years, as more years would make the data too large to code the patent filings manually.

B. Shapiro-Wilk Test

Table 6

Shapiro-Wilk Test of Normality

		Shapiro-Wilk		
Variable	Groups	W	df	P
DistanceLOG	Actual	.920	36	.012
	Random	.964	42	.201

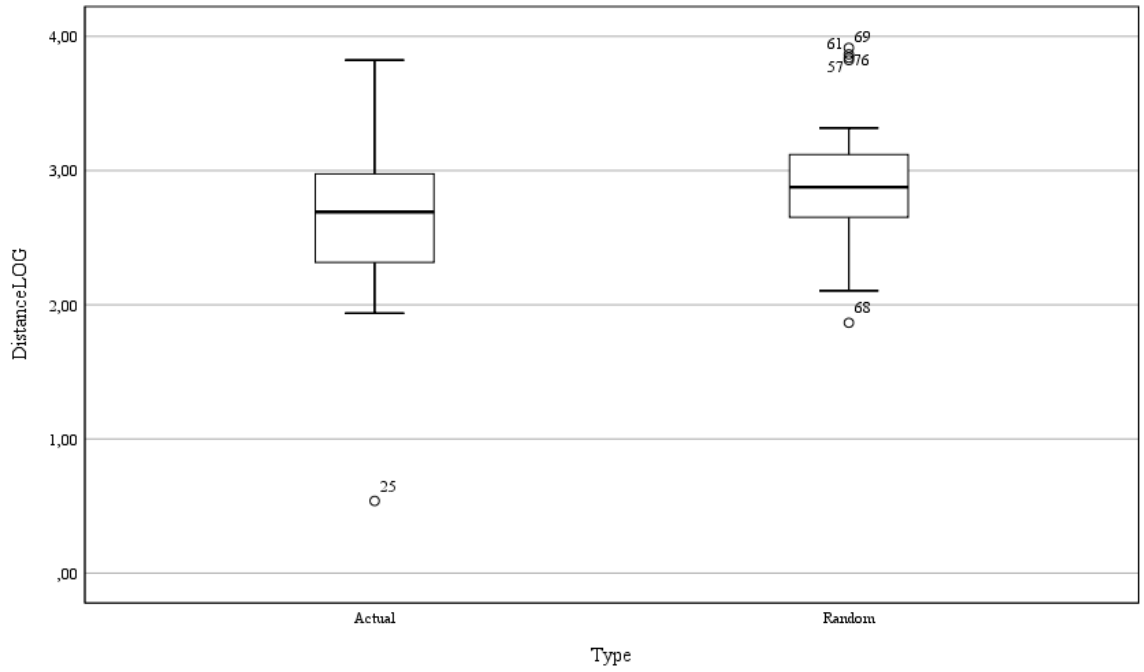

Figure 6. Boxplot of the variables in the actual and randomly created collaborations. The interquartile range of the actual collaboration is larger, confirming the result that the distribution does not follow a normal distribution.